

METHOD, SYSTEM, AND PROGRAM FOR
DUAL AGENT PROCESSES AND DUAL ACTIVE SERVER PROCESSES

BACKGROUND OF THE INVENTION

5 1. Field of the Invention

[0001] The present invention is related generally to dual agent processes coupled to dual active server processes.

2. Description of the Related Art

10 [0002] Some systems include many components (e.g., server computers executing server processes) and attempt to provide fault tolerance so that there is no single point of failure. For example, if one component in the system fails, the entire system does not fail.

[0003] Continual availability of data services is important for disaster recovery situations in which customers need to recover access to their data immediately. To have continual
15 availability of services provided, redundancy is important. Some conventional architectures include a primary server computer executing a primary server process and a backup server computer executing a backup server process. If an error occurs at the primary server process, the backup server process is used for processing while the primary server process is undergoing repair and/or recovery. The primary and backup server
20 processes may be located at different physical sites. The use of the backup server process at an alternate or remote site can reduce the recovery time of the primary server process because the backup server process is continually able to process new requests.

[0004] An agent process is software that acts on behalf of, for example, a user or server process, to perform tasks within an application program. In some conventional computer
25 systems, the agent process registers with the primary server process (i.e., notifies the server process that it exists). Then, the primary server process and the agent process can submit requests to each other, provide data to each other, etc. If the primary server process fails, then a user notifies the agent process that the agent process is to re-register

with the backup server process. For example, the user may use a Web browser at a console connected to the agent process to notify the agent process to re-register with another server process. The term "Web" refers to the World Wide Web, which is a network of server computers that contain programs and files, which in turn may contain
5 hypertext links to other documents available through the network. The term "Web browser" refers to a program that initiates requests to a Web server computer and displays the information that the Web server computer returns. The agent process then registers with the backup server process.

[0005] Requiring user intervention when a primary server computer and/or primary server
10 process fails is time consuming and inefficient. This may also lead to human errors. Therefore, there is a need in the art for improved use of dual server processes.

SUMMARY OF THE INVENTION

[0006] Provided are a method, system, and program for task processing and monitoring of
15 configuration and state information. A first server process and a second server process are detected. A first agent process is registered with the first server process. A second agent process is registered with the second server process. When a task is to be executed by the first server process, the task is executed with the first agent process. When the task is to be executed by the second server process, the task is executed with the second agent
20 process.

[0007] The described implementations of the invention provide a method, system, and program for providing two agent processes in a cluster, with each agent process registered with a different active server process, so that the failure of one agent/server pair does not impact the other agent/server pair. Additionally, a cluster may access storage devices, and
25 the configuration and state information of the storage devices of the cluster is sent to the active server processes via corresponding agent processes in the agent/server pairs.

BRIEF DESCRIPTION OF THE DRAWINGS

Referring now to the drawings in which like reference numbers represent corresponding parts throughout:

FIG. 1A illustrates, in a block diagram, a computing environment in accordance
5 with certain implementations of the invention.

FIG. 1B illustrates, in a block diagram, a computing environment in accordance with certain other implementations of the invention.

FIG. 1C illustrates, in a block diagram, further details of a computing environment illustrated in FIG. 1A in accordance with certain implementations of the invention.

10 FIG. 1D illustrates, in a block diagram, a server system architecture in accordance with certain implementations of the invention.

FIG. 1E illustrates flow of control between processes in accordance with certain implementations of the invention.

15 FIGS. 2A and 2C illustrate logic for processing initial configuration and state information at one cluster in accordance with certain implementations of the invention, while FIG. 2B illustrates logic for startup of server and agent processes at one cluster in accordance with certain implementations of the invention.

FIG. 3 illustrates logic for configuration and/or state changes at one cluster in accordance with certain implementations of the invention.

20 FIG. 4 illustrates task execution and response at one cluster in accordance with certain implementations of the invention.

FIG. 5 illustrates one implementation of the architecture of server systems in accordance with certain implementations of the invention.

25 DETAILED DESCRIPTION

[0008] In the following description, reference is made to the accompanying drawings which form a part hereof and which illustrate several implementations of the present invention. It is understood that other implementations may be utilized and structural and

operational changes may be made without departing from the scope of the present invention.

[0009] FIG. 1A illustrates, in a block diagram, a computing environment in accordance with certain implementations of the invention. Server systems 102 and 104 are
5 interconnected through a network 150, such as a local area network (LAN), wide area network (WAN), or the Internet. The Internet is a world-wide collection of connected computer networks (i.e., a network of networks). Each server system 102 and 104 has two clusters (which may also be referred to as "server system clusters"). Although each server system 102 and 104 are illustrated with two clusters, in alternate implementations
10 of the invention, each server system 102 and 104 may include fewer or more clusters. For example, server system 102 comprises clusters 111 and 112, which may be referred to as a left cluster 111 and a right cluster 112. Server system 104 comprises clusters 113 and 114. A cluster includes zero or more server processes and one or more pairs of agent processes. For example, cluster 111 includes agent processes (Agent A and Agent B), and
15 cluster 113 includes a server process (Server-A) and agent processes (Agent E and Agent F). In this illustration the server process executes on one of the clusters.

[0010] FIG. 1B illustrates, in a block diagram, a computing environment in accordance with certain other implementations of the invention. Server system 102 is connected to a host system 186 via a network 152, such as a local area network (LAN), wide area
20 network (WAN), or the Internet. In FIG. 1B, the server process 186 executes at the host system 188, rather than at a cluster. A host system may comprise any computing device known in the art, such as a mainframe, server, personal computer, workstation, laptop, handheld computer, telephony device, network appliance, virtualization device, storage controller, etc.

25 [0011] FIG. 1C illustrates, in a block diagram, further details of a computing environment illustrated in FIG. 1A in accordance with certain implementations of the invention. A networked domain 100 includes various server systems 102, 104, 106, and 108 interconnected through a network such as a local area network (LAN), wide area network

(WAN), or the Internet. The Internet is a world-wide collection of connected computer networks (i.e., a network of networks). A domain is a part of a computer network in which the data processing resources are under common control.

[0012] The server systems 102, 104, 106, and 108 in the network domain 100 may be
5 viewed as being located at Site A and Site B. In the illustrated example, Site A includes server systems 102 and 104, while Site B includes server systems 106 and 108. In certain implementations of the invention, the server systems 102 and 140 at Site A and the server systems 106 and 108 at Site B may be physically far apart. Although four server systems 102, 104, 106, 108 are illustrated, fewer or more server systems may be used without
10 departing from the scope of the invention. In certain implementations of the invention, server systems are implemented using an IBM TotalStorage Enterprise Storage Server™ (ESS), available from International Business Machines Corporation.

[0013] In addition to the clusters illustrated in FIG. 1A, FIG. 1C illustrates that server system 106 has clusters 115 and 116, and server system 108 has clusters 117 and 118.
15 When the network domain 100 is configured, some clusters may be configured to be "domain servers". Domain servers include server processes, while clusters that are not configured as domain servers do not include server processes. In the example of FIG. 1C, clusters 113 and 117 have been configured as domain servers. Also, the server system has storage devices that are shared by the clusters in that server system. For example, clusters
20 111 and 112 of server system 102 share storage devices 127.

[0014] Each cluster includes zero or more server processes and one or more pairs of agent processes. For example, cluster 111 includes two agent processes (Agent A 120 and Agent B 122 and no server processes, while cluster 113 includes a server process (Server A 130), as well as agent processes (Agent E 170 and Agent F 172). A cluster also
25 includes a driver process (e.g., Driver-1 124) and persistent data (e.g., Persistent Data-1 126) that is accessed by the agent processes in the cluster. Persistent data is stored in, for example, disk. When a cluster includes two agent processes, the agent processes may be referred to as a first agent process and a second agent process, and each agent process

registers with a different server process. A hardware console 160 includes a Web browser 162 that can connect to the server processes (Server A 130 and Server B 140) and allows, for example, a system administrator or other user to submit tasks to the server processes for execution.

5 [0015] Each cluster has two agent processes that may be invoked to process tasks. Agent processes are invoked by server processes, such as Server A 130 or Server B 140. Each server process is aware of all agent processes that have registered with that server process. In each cluster 111, 112, 113, 114, 115, 116, 117, and 118, one agent process registers with one server process, and the other agent process registers with the other server
10 process. For example, Agent A, Agent C, Agent E, Agent G, Agent I, Agent K, Agent M, and Agent O register with Server A 130. Agent B, Agent D, Agent F, Agent H, Agent J, Agent L, Agent N, and Agent P register with Server B 140. Then, Server A 130 and Server B 140 may invoke the respective registered agent processes to perform tasks. An agent process and a corresponding server process may be said to form an agent/server
15 application or agent/server pair. For example, Agent B 122 and Server B 140 may be said to form an agent/server application or agent/server pair.

[0016] In particular, when the network domain 100 is configured, two agent processes are set up per cluster. Each cluster is aware of each server process. Each server process has a list of all agent processes within the domain that are registered with that server process.
20 Each agent process in the cluster communicates with a different server process. Since each cluster includes two agent processes registered with different server processes, if one agent process fails, the other agent process may continue processing. If one server process in the domain 100 fails, the other server process may continue processing. Moreover, since a server system 102, 104, 106, 108 has two clusters, if one cluster fails,
25 the other cluster in the server system may continue processing.

[0017] FIG. 1D illustrates, in a block diagram, a server system 180 architecture in accordance with certain implementations of the invention. In FIG. 1B, each cluster 180

and 182 includes a server process (i.e., Server X and Server Y). The server system 180 may be part of a network domain (not shown).

[0018] FIG. 1E illustrates flow of control between processes in accordance with certain implementations of the invention. A Web browser 190 (e.g., Web browser 162) sends a task to a server process 192 (e.g., Server A 130). The server process 192 invokes an agent process 194 (e.g., Agent A 120) to perform the task. The agent process 194 forwards the task to a driver process 196 (e.g., Driver-1 124) for execution of the task. The driver process 196 executes the task and transmits information regarding task execution to the agent process 194. The agent process 194 transmits information regarding task execution to the server process 192. The server process 192 transmits information regarding task execution to the Web browser 190. Although examples herein refer to a particular process (e.g., a driver process) performing processing, the process may actually invoke one or more other processes to perform the processing.

[0019] FIGs. 2A and 2C illustrate logic for processing initial configuration and state information at one cluster in accordance with certain implementations of the invention, while FIG. 2B illustrates logic for startup of server and agent processes at one cluster in accordance with certain implementations of the invention. Control begins in FIG. 2A at block 200 with the initialization process launching. There are various conditions that may result in the initialization process launching. For example, the initialization process may launch when a server system is started (e.g., powered-on), when a user initiates the process using the Web browser 162, or when a system administrator initiates the process using the Web browser 162 or a different user interface.

[0020] In block 202, a collection and storage process that is automatically launched when the initialization process is launched collects initial configuration information (e.g., how many storage devices are in the cluster) and state information (e.g., whether each storage device is available or unavailable) and stores the initial configuration and state information as persistent data at the cluster.

[0021] In block 204 (FIG. 2B), it is determined whether the cluster is configured as a domain server. A cluster may be configured as a domain server by, for example, a system administrator, when the domain is being configured. When a cluster is configured as a domain server, the cluster includes a server process. In the example of FIG. 1C, cluster
5 111 is not configured as a domain server, but cluster 113 is configured as a domain server. Therefore, when the logic of FIGs. 2A, 2B, and 2C is performed for cluster 111, in block 204 it will be determined that cluster 111 is not configured as a domain server. When the logic of FIGs. 2A, 2B, and 2C is performed for cluster 113, it will be determined that cluster 113 is configured as a domain server.

10 [0022] If the cluster is configured as a domain server, processing continues to block 206, otherwise, processing continues to blocks 210 and 220. In block 206, the server process is launched (i.e., the server process is started). The server process is then able to process tasks received from the Web browser 162.

[0023] In certain implementations, processing at blocks 210-214 and 220-224 is
15 performed concurrently. In certain alternative implementations, processing at blocks 210-214 and 220-224 may be performed sequentially or some processing may be performed sequentially and other processing may be performed concurrently.

[0024] In block 210, it is determined whether a first server is configured. The first server is a domain server with which the first agent process is to register. If so, processing
20 continues to block 212, otherwise, processing is done. In block 212, a first agent process (e.g., Agent A 120) is launched. In block 214, the first agent process (e.g., Agent A 120) registers with the first server process (e.g., Server A 120). In block 215, it is determined whether the registration was successful. If so, processing continues to block 230, otherwise, processing returns to block 214 to attempt registration again. In certain
25 implementations of the invention, registration is attempted until the first agent process is successfully registered. In certain other implementations of the invention, registration is attempted a certain number of times (e.g., three times).

[0025] In block 220, it is determined whether a second server (e.g., Server B 140) is configured. The second server is a domain server with which the second agent process is to register. If so, processing continues to block 222, otherwise, processing is done. In block 222, a second agent process (e.g., Agent B) is launched. In block 224, the second agent process (e.g., Agent B) registers with the second server process (e.g., Server B). In block 225, it is determined whether the registration was successful. If so, processing continues to block 230, otherwise, processing returns to block 224 to attempt registration again. In certain implementations of the invention, registration is attempted until the first agent process is successfully registered. In certain other implementations of the invention, registration is attempted a certain number of times (e.g., three times).

[0026] In FIG. 2C, in block 230, active agent processes retrieve initial configuration and state information from persistent data in a cluster in which the active agent is located. An active agent is any agent that is running. In block 232, active agent processes send initial configuration and state information to corresponding server processes. For example, the first agent process (e.g., Agent A) sends initial configuration and state information to the first server process (e.g., Server A), and the second agent process (e.g., Agent B) sends initial configuration and state information to the second server process (e.g., Server B). Each server process maintains a copy of the data sent by the agent processes. For example, Server A 130 has a copy of data 132, and Server B 140 has a copy of data 142.

[0027] FIG. 3 illustrates logic for configuration and/or state changes at one cluster in accordance with certain implementations of the invention. Control begins at block 300 with the driver process at the cluster (e.g., Driver-1 124 in cluster 111) receiving changed configuration information (e.g., a storage device has been added or deleted from the cluster) and/or state information (e.g., a storage device has gone from available to unavailable). Although examples herein refer to a particular process (e.g., a driver process) performing processing, the process referred to may actually invoke one or more other functions or processes to perform the processing.

[0028] In block 302, the changed configuration and/or state information is stored as persistent data at the cluster. In block 304, the active agent processes are sent messages by the driver process indicating that there has been a change in configuration and/or state at the cluster. In certain implementations of the invention, when the driver process
5 recognizes a change in configuration and/or state information, the driver process broadcasts the information (e.g., the addition or removal of a resource) within the cluster. In block 306, active agent processes retrieve changed configuration and/or state information from persistent data. In block 308, the active agent processes send changed configuration and/or state information to corresponding server processes.

10 [0029] FIG. 4 illustrates task execution and response at one cluster in accordance with certain implementations of the invention. Control begins at block 400 with a Web browser issuing a task to a first server process (e.g., Server A). In certain implementations of the invention, a user selects a server process that is to execute a task at the hardware console 160. In block 402, the first server process sends the task to a first agent process. In block
15 404, the first agent process stores its own identification (e.g., "Agent A") as persistent data and sends the task to a driver process. In block 406, the driver process executes the task.

[0030] In block 408, it is determined whether the task failed. If so, processing continues to block 410, otherwise, processing continues to block 412. In block 410, task failure
20 information is stored as persistent data in the appropriate cluster of the agent process. In block 412, the driver process retrieves the invoking agent process identification from persistent data, which in this case is the first agent process identification, and notifies the first agent process about task completion. In certain implementations of the invention, the driver process invokes a task completion process to update task failure information in the
25 persistent data, determine which agent process requested task processing, and to notify the agent process of the task completion.

[0031] In block 414, it is determined whether the task failed. If so, processing continues to block 416, otherwise, processing continues to block 418. In block 416, the first agent

process retrieves task failure information from persistent data. In block 418, the first agent process sends task completion status to the first server process. In block 420, the first server process sends task completion status and failure information, if there has been a task failure, to the browser. This logic occurs for each other server process (e.g., Server
5 B) and agent process (e.g., Agent B) pair.

[0032] In summary, the configuration of implementations of the invention allows for two or more fully functional servers to be configured for all clusters in the domain. Both server processes in the domain are detected when a cluster is started. The server process in the domain is started when a cluster is started that is configured as a domain server.
10 Also, two agent processes in the cluster are started, one for each server process. Each agent process manages communication with a corresponding server process.

[0033] If an error occurs between one agent/server pair and the pair fails, another agent/server pair would be unaffected by this failure. Thus, if a task fails on one agent/server pair due to unavailability of the pair, a user can tell a second agent/server pair
15 to execute the task. With implementations of the invention, since there are two agent processes, the communication is performed in parallel, concurrently.

[0034] Furthermore, the techniques described herein for dual active server processes are applicable in any network environment in which there are some set of agent processes that need to be coordinated. Implementations of the invention provide the ability of agent
20 processes to interoperate with a pair of server processes in a way that provides continuous availability to a server process when at least one of the dual active servers is active. The underlying processes that are being executed by the server are not limited to any specific application.

[0035] Although examples herein refer to monitoring of configuration and state
25 information of storage devices, implementations of the invention are not limited to such monitoring. That is, the dual active server processes described in the invention are easily extended by those skilled in the art to other applications besides the monitoring of storage devices. In particular, implementations of the invention are applicable to any set of

processes or services that can be requested by multiple sources (e.g., agent processes) and are managed by a one or more providers (e.g., server processes) to provide non-disruptive operation of an overall system. Examples of other applications that might benefit from the invention include, but are not limited to, real time management and control of physical
5 plants (factories, power plants, machinery), real time management of online transaction processing systems (Automated Teller Machines (ATMs), stock exchanges, airline ticketing, etc.), real time management of computer systems (disaster recovery processes, networked printer queues, etc.), and real time management of communication networks (telephony, private branch exchanges (PBXs), etc.).

10 [0036] Enterprise Storage Server is a registered trademark or trademark of International Business Machines Corporation in the United States and/or other countries.

Additional Implementation Details

[0037] The described techniques may be implemented as a method, apparatus or article of
15 manufacture using standard programming and/or engineering techniques to produce software, firmware, hardware, or any combination thereof. The term "article of manufacture" as used herein refers to code or logic implemented in hardware logic (e.g., an integrated circuit chip, Programmable Gate Array (PGA), Application Specific Integrated Circuit (ASIC), etc.) or a computer readable medium, such as magnetic storage
20 medium (e.g., hard disk drives, floppy disks, tape, etc.), optical storage (CD-ROMs, optical disks, etc.), volatile and non-volatile memory devices (e.g., EEPROMs, ROMs, PROMs, RAMs, DRAMs, SRAMs, firmware, programmable logic, etc.). Code in the computer readable medium is accessed and executed by a processor. The code in which described embodiments are implemented may further be accessible through a transmission
25 medium or from a file server over a network. In such cases, the article of manufacture in which the code is implemented may comprise a transmission media, such as a network transmission line, wireless transmission media, signals propagating through space, radio waves, infrared signals, etc. Thus, the "article of manufacture" may comprise the medium

in which the code is embodied. Additionally, the "article of manufacture" may comprise a combination of hardware and software components in which the code is embodied, processed, and executed. Of course, those skilled in the art will recognize that many modifications may be made to this configuration without departing from the scope of the present invention, and that the article of manufacture may comprise any information bearing medium known in the art.

[0038] The logic of FIGs. 2A, 2B, 2C, 3, and 4 describe specific operations occurring in a particular order. In alternative implementations, certain of the logic operations may be performed in a different order, modified or removed. Moreover, steps may be added to the above described logic and still conform to the described implementations. Further, operations described herein may occur sequentially or certain operations may be processed in parallel, or operations described as performed by a single process may be performed by distributed processes.

[0039] The illustrated logic of FIGs. 2A, 2B, 2C, 3, and 4 may be implemented in software, hardware, in programmable and non-programmable gate array logic or some combination of these.

[0040] FIG. 5 illustrates one implementation of the architecture of the server systems 102, 104, 106, and 108 in accordance with certain implementations of the invention. The computer systems 102, 104, 106, and 108 may implement a computer architecture 500 having a processor 502 (e.g., a microprocessor), a memory 504 (e.g., a volatile memory device), a network card 508, and storage 510 (e.g., a non-volatile storage area, such as magnetic disk drives, optical disk drives, a tape drive, non-volatile RAM, etc.). An operating system 505 may execute in memory 504. The storage 508 may comprise an internal storage device or an attached or network accessible storage. Computer programs in the storage 510 are loaded into the memory 504 and executed by the processor 502 in a manner known in the art. A network card 508 enables communication with a network. An input device 512 is used to provide user input to the processor 502, and may include a keyboard, mouse, pen-stylus, microphone, touch sensitive display screen, or any other

activation or input mechanism known in the art. An output device 514 is capable of rendering information transmitted from the processor 502, or other component, such as a display monitor, printer, storage, etc.

[0041] The computer architecture 500 may comprise any computing device known in the art, such as a mainframe, server, personal computer, workstation, laptop, handheld computer, telephony device, network appliance, virtualization device, storage controller, etc. Any processor 502 and operating system 505 known in the art may be used.

[0042] The foregoing description of implementations of the invention has been presented for the purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed. Many modifications and variations are possible in light of the above teaching. It is intended that the scope of the invention be limited not by this detailed description, but rather by the claims appended hereto. The above specification, examples and data provide a complete description of the manufacture and use of the composition of the invention. Since many implementations of the invention can be made without departing from the spirit and scope of the invention, the invention resides in the claims hereinafter appended.